Network Working Group
Request for Comments: 1092

J. Rekhter
T. J. Watson Research Center
February 1989

EGP and Policy Based Routing in the New NSFNET Backbone

Status of this Memo

This memo discusses implementation decisions for routing issues in the NSFNET, especially in the NSFNET Backbone. Of special concern is the restriction of routing information to advertize the best route as established by a policy decision. Distribution of this memo is unlimited.

Introduction

The NSFNET backbone routes packets between the Regionals Networks to which it is connected, (i.e., the packets arriving at a backbone entry node are routed to an exit node). How they travel through the network is determined by two components:

the NSFNET backbone routing protocol/algorithm, and

additional information about the externally connected networks.

This paper is concerned with how reachability information between the external networks and the NSFNET backbone is exchanged so that packets can be routed to the correct destination by using a reasonable path.

EGP as reachability protocol

The EGP (Exterior Gateway Protocol) routing method will be used to exchange reachability information between the NSFNET backbone and the regional networks.

There are several problems with using EGP as a reachability protocol for routing in a meshed environment. Some EGP components require further definitions for the NSFNET backbone - regional network interactions. It should be noted that the use of EGP is only viewed as an interim measure until better inter autonomous system protocols are defined and widely deployed for gateways used by regional networks.

The following is a list of some EGP problems and issues:

The EGP model assumes an engineered spanning tree topology,

however, the NSFNET (due to the presence of backdoor routes) does not fit into this model. In the NSFNET the same network may be advertized as reachable by more than one regional network. Besides the fact that the overall NSFNET does not fit into a spanning tree model there are serious concerns with the concept of the "core" (central to the EGP) and its obvious deficiencies.

While EGP is going to isolate intra-Regional routing from the intra-NSFNET-Backbone routing, it does not address the issue of false information which may be supplied by regional networks. EGP by itself does not protect a particular network from unwanted and unsolicited representation by some regional network. As an example, if network N1 is reachable through regional network R1 as well as through regional network R2, EGP has no provisions to specify one of these paths as a primary and one as a secondary, since there is not generally accepted interpretation of EGP metrics today. Also, there is nothing in EGP which can prevent one or more regional networks from advertizing other networks (in particular, networks which belong to other regional networks) as reachable with zero distance. This could result in the creation of a "black hole" or at least in suboptimal IP routing.

EGP by itself has no provisions to guarantee that routes through the NSFNET Backbone will be preferred over routes through the backdoor routers or vice versa.

Policy Based Routing

Looking at the problems listed above the appearance of the new factors like autonomy and mutual trust becomes obvious. While trying to achieve the routing functionality required for the new NSFNET backbone we should realize that one of our primary concerns has to be the accommodation of those new factors.

This means that some kind of a rudimentary Policy Based Routing method becomes imperative. We would like to emphasize, however, that we are not talking about complete Policy Based Routing, but that we are rather concerned about supporting a minimum subset of a policy functionality to be an initial solution to the above mentioned problems. This requires support and cooperation between the management of each of the networks connected to the NSFNET backbone.

We need to support the ability of a particular network N, which belongs to one of the regional networks, to establish a bilateral agreement with one or more regional networks of the type "network N can be reached via one or more regional networks (RN1, RN2, ... RNx)". This allows each network to select one or more representatives at the regional network level. Once this agreement

is established the information will be available to:

The network which initiated the agreement.

The management of the regional network(s) with whom this agreement has been established.

The NSFNET backbone Network Operation Center where it will be entered into the Routing Policy Data Base which will be available through the NSFNET information services.

Supporting multiple routes to the NSFNET core requires the guarantee that for a certain network N, no regional network other than the one(s) selected by N, will advertize N as reachable, which necessitates that the NSFNET core will ignore unauthorized advertisements for network N.

EGP and Rudimentary Policy Based Routing

Each network which belongs to the NSFNET will select a specific regional network as its primary representative to the NSFNET core by bilateral agreement with the management of same regional network as well as the NSFNET backbone management. The same network can furthermore select an arbitrary number of other regional networks as their secondary, tertiary, etc., representative by establishing bilateral agreements with the management of the corresponding regional networks as well as the NSFNET backbone management.

Reachability information supplied by each regional network will be distributed to all other NSS nodes of the NSFNET Backbone. We would like to emphasize that we are not going to flood EGP packets internally within the backbone, but to rather use the learned information for the interior gateway protocol, which uses the ANSI IS-IS protocol.

The implementation allows for a defined regional network to advertize a particular leaf network in the EGP NR packets with a distance of zero. Secondary representatives may advertize the same network with distance one or higher. If the path through the primary regional representative is available all secondary paths will be ignored. If the path through the primary regional representative goes down (which will be discovered via the EGP NR information), the next path with the lowest available EGP metric will be used.

We will also be able to detect and report unsolicited representations. This will be done by examining (on a periodic basis) all reachability information obtained via EGP. The result will be compared against the Routing Policy Data Base which will hold

information about all bilateral agreements between networks and their regional representatives. Any mismatch will cause an alarm to the Network Operations Center. For example, network N established a bilateral agreement with the regional network R1 electing it as its primary representative. The EGP NR record received from the regional network R5 advertizes the network N as reachable with distance zero. By comparing the Routing Policy Data Base entry for the network N with the EGP NR record a mismatch will be detected and an alarm is forwarded to the Network Operation Center.

Since the whole scheme is based on a combination of the network number and the autonomous system number, to allow for further verification, it is also important to insure the correctness of the autonomous system numbers as advertized by the regionals networks to the NSFNET core.

The autonomous system number validation for each regional network will be performed at the NSS which connects the particular leaf network to the NSFNET backbone. All discrepancies wil be reported to the Network Operations Center.

The NSFNET backbone will be considered as a separate Autonomous System with its own autonomous system number.

Backbone versus Backdoor Routes

There are instances where regional networks prefer paths through some backdoor route over paths through the NSFNET backbone. Therefore, the reachability information advertized by the NSFNET core to the regional networks (via EGP NR records) will always use a fixed metric of 128 for all routes. This may aid to encourage traffic to flow through backdoors, if desired and available.

The regional networks can use a variety of techniques to determine how they route traffic for any particular network at their own option.

What do we expect from the Regional Networks

Each regional network should get its own Autonomous System number.

The connection between regional networks to NSFNET backbone will be done via EGP. It is the responsibility of the regional backbone to provide an EGP functionality via the attachment to the E-PSP dedicated to the regional network.

The EGP functionality may require a translation of network numbers in and out of the regional network. In any case, the NSFNET backbone

expects individual network numbers of the leaf networks of the regional network, as long as they should be advertised, and will announce individual networks known to the NSFNET core to the regional network.

The EGP support should includes the ability to configure EGP metrics from some statically definable configuration table. If the EGP metrics cannot be defined or if they are not fixed the metric determination will be done by the NSFNET backbone routers, as taken from their databases, themselves. In that case, it is the responsibility of the regional network to provide the NSFNET backbone management with the metric data to allow for proper use of metrics.

We also expect each regional network to handle all bilateral agreements with its leaf networks regarding Policy Based Routing and supply a copy of those agreements to the NSFNET backbone management.

Acknowledgements

I would like to express my thanks to Barry Appelman (T.J. Watson Research Center, IBM Corp.) and Hans-Werner Braun (Merit) for their contributions to this document.

Author's Address

Jacob Rekhter
T.J. Watson Research Center
IBM Corporation
P.O. Box 218
Yorktown Heights, NY 10598

Phone: (914) 945-3896

Email: YAKOV@IBM.COM